Population studies of the fragile X: a molecular approach

P A Jacobs, H Bullman, J Macpherson, S Youings, V Rooney, A Watson, N R Dennis

Abstract
The fragile X mutation can now be recognised by a variety of molecular techniques. We report a pilot screening survey of a population of children with mental impairment in which we used Southern blotting methods to detect the fragile X mutation, augmented by cytogenetic studies on children whose phenotype suggested a possible chromosome abnormality. There were 873 children with special educational needs in our survey and 310 fulfilled our criteria for testing. A sample was obtained from 254, of whom four were found to have a full fra(X) mutation (ΔL) and none to have a premutation.

The number of CGG repeats in our population of X chromosomes was measured by PCR analysis and the genotype at the closely linked polymorphic locus FRAXAC1 established. The distribution of CGG repeat numbers was very similar to that of the control population reported by Fu et al and the distribution of FRAXAC1 alleles almost identical to that of the control population reported by Richards et al. Among the non-fragile X chromosomes, we found a very significant correlation between the size of the CGG repeat and the FRAXAC1 genotype. There was a dearth of A and D genotypes in subjects with a small number of CGG repeats and an excess of the A genotype in those with a large number of CGG repeats. These observations are considered in the light of the reported disequilibrium between the A (and possibly also the D) genotype and the fra(X) mutation.

(fragile X) mutation is associated with a cytogenetically detectable fragile site, after which the syndrome is named. The fragile site is only seen when the cells of affected subjects are grown under conditions of thymidine perturbation. The cytogenetic marker is a good test for retarded males and most, but not all, retarded females; however, it fails to detect a proportion of retarded females and virtually all non-retarded carriers of both sexes. In spite of these disadvantages, until the advent of molecular diagnosis, the cytogenetic detection of the fragile site at Xq27.3 was the only test for the syndrome.

Recent research has shown that the fra(X) mutation results from expansion of a CGG repeat in the 5' untranslated region of the FMR-1 gene. The number of CGG repeats differs among people, but is commonly between five and 50 in the normal population. Unaffected carriers of the fra(X) mutation have an increased number of repeats, usually between 60 and 200, and are said to have a premutation or a small expansion (ΔS), while retarded subjects and some non-retarded carrier females have over 200 repeats and are said to have a full mutation or large expansion (ΔL). The full mutation is associated with methylation and failure of expression of the FMR-1 gene (table 1). No subject with either a small or large expansion has yet been found to be a de novo mutant, all having a parent who themselves has a small or large expansion. There is a tendency in families in which the fra(X) mutation is segregating for there to be an increase in size of the expansion from generation to generation when inheritance is through a female, while there is little alteration in size when the inheritance is through a male.

Preliminary evidence suggests that (1) repeats numbering fewer than 40 are stable within a person and are stably inherited, that is, they are both mitotically and meiotically stable; (2) repeat numbers from 40 to 60 are stable within a person and may be stably

Table 1  Fragile X expansion.

<table>
<thead>
<tr>
<th>PCR analysis (No of CGG repeats)</th>
<th>Southern analysis (increase in bp = Δ)</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>~5-50</td>
<td>Mitotic</td>
</tr>
<tr>
<td>Premutation</td>
<td>~40-200</td>
<td>Stable</td>
</tr>
<tr>
<td>Full mutation</td>
<td>&gt;200</td>
<td>Unstable</td>
</tr>
</tbody>
</table>

Wessex Regional Genetics Laboratory, Salisbury District Hospital, Odstock, Salisbury SP2 8BJ, UK.
P A Jacobs
H Bullman
J Macpherson
S Youings

Community Child Health Department, Odstock Hospital, Salisbury SP2 8BJ, UK.
V Rooney

Department of Community Paediatrics, Newport, Isle of Wight, UK.
A Watson

Wessex Clinical Genetics Service, Princess Anne Hospital, Cowford Road, Southamptom SO9 4HA, UK.

Correspondence to Dr Jacobs.

Received 9 February 1993.
Revised version accepted 23 March 1993.
Population of the fragile X: a molecular approach

455

Table 2 Demographic data.

<table>
<thead>
<tr>
<th>Population of 5–18 year olds</th>
<th>35 675</th>
</tr>
</thead>
<tbody>
<tr>
<td>Records of SEN pupils examined</td>
<td>8 773</td>
</tr>
<tr>
<td>Eligible children</td>
<td>3 103</td>
</tr>
</tbody>
</table>

Materials and methods

The study population consisted of children aged 5 to 18 years with SEN attending schools in the Salisbury and Isle of Wight Health Authority areas. The community health records from this group were reviewed and any who had documented medical reasons for their special needs (for example, Down's syndrome, cerebral palsy) were discarded. Those selected for testing were children for whom special educational provision was required, predominantly because of intellectual disability, and in whom no medical diagnosis for their special needs was noted. Details of the study population are given in tables 2 and 3.

A letter was sent to the parents of the children selected for testing giving a short explanation of fragile X and requesting permission to give their child a brief physical examination and take a sample of venous blood. A form to sign and a stamped addressed envelope were enclosed for their reply. If no reply was received from the parents after two weeks, a reminder was sent.

Children were seen in their schools as this caused minimum disruption and only involved the parents if they wished to be present. A brief clinical examination without undressing the child was carried out before each child was bled. The face, ears, palate, teeth, and hands were examined and any unusual features noted. Head circumference and ear length were measured. A 5 to 10 ml sample of blood was taken into EDTA in all cases and, if a chromosome anomaly was suspected on clinical grounds, a sample was also taken into lithium heparin.

If the results were normal the parents were informed by letter. If fragile X or a chromosome abnormality was found, a clinical geneticist contacted the parents directly. A repeat sample was taken by DNA and chromosome testing from each subject who had an abnormal result. When the result was confirmed the family was seen by a clinical geneticist and blood obtained from other appropriate family members. The family doctors were informed of the result in all cases.

Table 3 Study population.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligible children</td>
<td>213</td>
<td>97</td>
<td>310</td>
</tr>
<tr>
<td>Declined test</td>
<td>15</td>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td>No reply received</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Unable to contact</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Unable to bleed</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Blood obtained</td>
<td>180</td>
<td>74</td>
<td>254*</td>
</tr>
</tbody>
</table>

* Includes eight where the child could not be bled but the parents were tested instead.

Inherited or may alter in size when transmitted from parent to offspring, that is, mitotically stable but behaviour at meiosis variable; (3) repeat numbers between 60 and 200 are stable in a person but not stable from generation to generation, that is, mitotically stable but meiotically unstable; and (4) repeat numbers over 200 are unstable among cells and tissue of the same subject and when inherited from parent to offspring, that is, both mitotically and meiotically unstable.

Recently two polymorphic CA repeat sequences, FRAXAC1 and FRAXAC2, which flank the fra(X) repeat have been described. These markers show no recombination with the fra(X) mutation or with one another. It has been suggested that these markers are not only in strong linkage disequilibrium with one another but also with the fra(X) mutation which is more frequent on certain haplotypes than on others. For example, three haplotypes which account for only 15% of the normal Caucasian population are found in 58% of fra(X) chromosomes. Furthermore Richards et al. suggested that this might be because certain haplotypes have a larger number of CGG repeats which might make them more mutable, and thus more likely to give rise to a premutation or full mutation.

A large number of surveys of mentally impaired males and females have been published in which the cytogenetic detection of the fragile site at Xq27.3 was used to ascertain probands with the fragile X syndrome. The most comprehensive of these surveys was that of Turner et al. who screened the retarded population of New South Wales. The advent of a molecular test for the fra(X) syndrome allows a different approach to population screening. We decided to carry out systematic screening of a population of children with special educational needs (SEN) using a series of molecular tests, augmented by cytogenetic testing of children whose phenotype suggested a cytogenetic abnormality. The initial molecular screen used Southern analysis to identify the AS and AL expansions, and was followed by PCR tests to determine where possible the exact CGG repeat number and also to type linked (CA)n markers.

The purpose of our study was fivefold: (1) to ascertain the efficacy of Southern blot analyses for population screening; (2) to ascertain the number of affected subjects (AL) detected by this technique; (3) to determine the number of CGG repeats on the X chromosomes by PCR technology from which the number of premutations may be inferred; (4) to determine the distribution of the FRAXAC1 polymorphism in our population; and (5) to determine the relationship between the FRAXAC1 genotype and the number of CGG repeats in our population of X chromosomes. We originally set out to determine the distribution of haplotypes at the fra(X) site by typing the bracketing polymorphisms FRAXAC1 and FRAXAC2; unfortunately, in our hands, the FRAXAC2 marker could not be typed reliably. However, as FRAXAC1 and FRAXAC2 are in strong
Table 4 Results: fragile X.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children tested</td>
<td>180</td>
<td>74</td>
<td>254</td>
</tr>
<tr>
<td>Fragile X positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number*</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>%</td>
<td>2.2</td>
<td>0</td>
<td>1.6</td>
</tr>
</tbody>
</table>

* Subjects showing a large (AL) DNA expansion with probe StB12.3.

LABORATORY METHODS

The status of the CGG repeat was measured in two ways: firstly on Southern blots of human genomic DNA digested with EcoRI and the methylation sensitive enzyme BstZI hybridised to the intragenic probe StB12.3; this enables DNA fragment sizes to be estimated and expansions measured by comparison with a set of ‘marker’ fragments of known size, and simultaneously determines the methylation status of the fragment. The procedures used for digestion, blotting, and labelling were as previously described.6–7 Secondly, we used PCR to determine accurately the number of CGG repeats present on the population of X chromosomes studied, applying the method described by Fu et al.8

Cytogenetic studies using standard protocols were carried out on 40 patients who were considered appropriate on clinical grounds. Cytogenetic studies in conditions suitable for detecting the fragile site at Xq27.3 were carried out on all probands who showed a ΔL on molecular testing.

The CA repeat marker FRAXAC1 was studied using the protocols described by Richards et al.4 but with standard PCR buffer (Promega) and 2 pmol of one primer end labelled with γ-[32P]dATP using T4 polynucleotide kinase (Cambio).

Results

A total of 310 children was selected for screening for the fragile X. However, 14 of these families could not be contacted and thus the study population consisted of 296 children. Of these, 257 (87%) agreed to the testing and a blood sample was obtained from 254 (86%). The reasons for the 42 ‘failures’ are given in table 3. Forty of the 257 children who were seen were also screened by conventional cytogenetic analysis.

Among the 180 males screened for the fragile X, using probe StB12.3 a total of four (2.2%) was found to carry a full (ΔL) mutation while no full mutations were seen among the 74 females tested. No premutations were detected in any of the children (table 4). One male was found with a female pattern of DNA analysis on exposure to StB12.3; that is, two bands indicating an active and inactive X chromosome, and subsequent chromosome analysis showed him to have a 47,XXX chromosome constitution. Chromosome analysis was possible on only three of the four males with full mutations and showed the fragile site to be present in substantial numbers of cells (12%, 30%, and 32%). One fra(X) proband was already known to us and his pedigree is shown in fig 1, pedigree A. He has four sibs one of whom is a mentally retarded brother with a ΔL; the mother who also has a ΔL is intellectually normal. One fra(X) proband’s family was living abroad and no further studies have yet been done. The two remaining families (fig 1, pedigrees B and C) are currently under investigation. Both boys had a mentally impaired sister with ΔL. Neither of these girls was included in our study because they were both beyond school age. As can be seen, the mutation in both families has been traced back to a maternal grandmother with a ΔS.

Among the 40 patients selected for cytogenetic analysis four were found to be abnormal; one male had a 46,XY,del(13)(pter→q13)de novo and one female a 46,XX,r(22) (p11q13.3)de novo constitution. Two brothers had a deleted Y chromosome but this was also seen in the clinically normal father and judged to be of no relevance to their SEN.

The number of CGG repeats was measured by PCR analysis in a total of 309 unrelated X chromosomes from our study and the results are shown in fig 2. As can be seen, excluding the four with ΔL, the numbers range from 13 to 49. The major peak extends from 28 to 31 with 22.5% of the X chromosomes having 29

Figure 1 Pedigrees of three of the four probands with fra(X).
repeats; there is also a minor peak between 19 and 23 which itself appears bimodal, with two subpeaks at 19 and at 22. There appears to be a relative dearth of X chromosomes with 21, 24, 26, or 27 repeats.

The results of typing using the closely linked CA repeat probe FRAXAC1 are shown in Table 5 and the distribution compared with that reported by Richards et al for 202 normal unrelated X chromosomes from CEPH families and Australian blood donors.

The relationship between CGG repeat number and the FRAXAC1 genotype was determined in the 160 unrelated non-fragile X males in our survey. Females were excluded because phase was unknown. Inspection of the histogram (Fig 2) suggested two antinodes, one at 26 and one at 33 repeats. These were used to divide the males into three classes: class 1 with fewer than 27 repeats, class 2 with 27 to 33 repeats, and class 3 with 34 or more repeats. Among the 160 males only three FRAXAC1 alleles, namely A, C, and D, were represented. The observed number of males in each CGG class is shown by FRAXAC1 genotype in Table 6 and compared with that expected if CGG size and FRAXAC1 were segregating independently. As can be seen, there is a highly significant deviation between observed and expected (χ² = 45.1, p < 0.001), with an absence of A and a reduction of D alleles in class 1 and a very marked excess of the A allele and a reduction in the numbers of C alleles in class 3. Sixteen females in our study had one X chromosome in which the number of CGG repeats was 34 and 10 had an A allele at the FRAXAC1 locus, many more than expected by chance.

**Discussion**

Our survey was clearly found to be acceptable both to the parents and teachers of the SEN children. We identified 873 SEN children in a population of 35 675 (2.4%) and detected four fra(X) males among the 180 tested and no fra(X) females among the 74 tested. Thus among the total population of 873 SEN children four (0.5%) were fra(X) positive. This figure is considerably lower than that of Turner et al who screened a total of 14 225 intellectually handicapped subjects of all ages in a population of 1.2 million (0.18%) and identified 253 (1.8%) fra(X) patients. As Turner et al screened a much smaller proportion of their population, their sample must have been more rigorously selected than ours and presumably relatively enriched for fra(X) subjects. In our survey none of the fra(X) children would have been missed and only 102 children (4%) would have been bled if we had only taken blood when head circumference and ear length were above the 50th centile for age. These two measurements are known to be increased in fra(X) children and Turner (personal communication) has shown that they can be used to make screening more selective with little or no reduction in detection of fra(X) subjects. If we assume that in both this survey and that of Turner et al all mentally impaired fra(X) were detected, we have a population frequency of 4 in 35 675 or 1 in approximately 9000 while Turner et al have a population frequency of 253 in 1.2 million or 1 in 5000. These are of course very minimal estimates and Turner et al found the prevalence rates among children in the public school system to be 1 in 2610 males and 1 in 4221 females.

No expansion that clearly fell into the premutation size range was detected among the 309 independent X chromosomes tested. Eleven were in the range 41 to 49, the expansion potential of which is at present unknown, although unstable alleles have been described with repeat numbers as low as 43. We have been unable to follow the pattern of inheritance of these larger normal alleles, as that was not part of our original protocol.

The distribution of normal allele size among the X chromosomes in our population is very similar to that published by Fu et al for 492 alleles obtained from three normal populations, namely mothers in CEPH families, blood donors from Texas, and normal spouses in North American fra(X) families. They, like us, have a major peak between 28 and 31 with approximately 30% of their X chromosomes having 29 repeats and a bimodal minor peak between 19 and 23. We appear to have a slightly wider spread with more alleles with 20

---

**Table 5** Frequency of FRAXAC1 alleles and comparison with published data.

<table>
<thead>
<tr>
<th>FRAXAC1 alleles*</th>
<th>Z</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No %</td>
<td>No %</td>
<td>No %</td>
<td>No %</td>
<td>No %</td>
<td>No %</td>
<td>No %</td>
<td>No %</td>
</tr>
<tr>
<td>Present study†</td>
<td>2 0.7</td>
<td>23 8</td>
<td>2 0.7</td>
<td>224 74</td>
<td>52 17</td>
<td>1 0</td>
<td>304</td>
</tr>
<tr>
<td>Richards et al</td>
<td>0 7</td>
<td>7 0</td>
<td>2 2</td>
<td>24 75</td>
<td>17 17</td>
<td>0 0</td>
<td>202</td>
</tr>
</tbody>
</table>

† Excluding the four AL patients who had A, A, C, and D alleles.

---

**Figure 2** Histogram of CGG sizes in 309 independent X chromosomes.
and 23 repeats than Fu et al.8 However, the close similarity of the data from the two sources indicates that, as one would have supposed, the distribution of CGG repeat copy number is the same in a mentally handicapped fra(X) negative population as in a ‘normal’ one and thus, with the exception of the subjects with ΔL, there is no correlation between the number of CGG repeats and mental ability. The absence of ΔS in our study is circumstantial evidence supporting the observation that carriers of the premutation have a normal distribution of IQ, and suggests that premutations, at least those with over 50 repeats, are unlikely to be present in more than 1% to 2% of the population. However, as the lower limit of premutation is at present undefined, and may be definable only in a statistical sense, there is an urgent need to determine the frequency and segregational behaviour of an unbiased series of alleles in the size range of the interface of normal and premutation.

The distribution of allele sizes of the CGG repeat, when accurately measured by PCR, in the normal population is quite remarkable and is similar in all ethnic groups studied to date (Warren, personal communication). In this respect it is different from the distribution of allele sizes in the trinucleotide repeats whose expansion is associated with myotonic dystrophy and Kennedy’s disease, as in both there is marked ethnic variation.10,11 The maintenance of such an unusual distribution of alleles across many populations argues against its origin by pure chance, and it is possible that CGG copy number even within the ‘normal’ range is subject to natural selection, though no convincing model has yet been proposed for such a circumstance.

The distribution of linked FRAXAC1 repeat alleles found in our 309 non-fra(X) chromosomes shows no significant difference from that reported from an Australian population (table 5) except for the appearance of one larger allele (Z) not found by Richards et al.1 Meanwhile, although only four fra(X) chromosomes were found in our survey, these do appear to show a similar excess of the A allele, compared to its rather low frequency on non-fra(X) chromosomes. This result, together with the data of Richards et al10 and our own unpublished observations on fragile X families, indicates that fragile X shows a positive linkage disequilibrium with FRAXAC1 allele A and an even more striking negative disequilibrium with allele C. Furthermore there is a very significant correlation between FRAXAC1 alleles and the number of CGG repeats in non-fragile X chromosomes, with an excess of the C allele among the chromosomes with the smallest number of repeats and an excess of the A allele among those chromosomes with the largest number of CGG repeats. This suggests that the number of CGG repeats on normal chromosomes may be related to their probability of expansion to premutation and full fragile X mutations. Indeed it has been shown12 that the probability of mutation in a microsatellite increases with the copy number. However, it is not clear whether the copy number is the only determinant of expansion of the trinucleotide repeat or whether other factors play a role.

One might suppose (1) that the A allele of FRAXAC1 is associated with a factor which predisposes to CGG expansion, or (2) that the C allele is associated with a factor which protects against CGG expansion, or (3) that a factor predisposing to CGG expansion is a single, or one of a very few mutations which occurred by chance on a less common background haplotype, that is, a ‘founder’ mutation. Resolution of this issue might be possible with the analysis of other closely linked polymorphisms in the region of fragile X, a comparison of the expansion potential of similar sized CGG alleles of different genotypes, or perhaps from DNA sequence data.

Clearly, there is still much to be learned about the origin of the fragile X mutation and it will need extensive population studies and possibly some quite complex models to account for all the accumulated observations. However, a substantial benefit to patients and their families will be gained if, through these studies, a consistent correlation can be found between the risk of fragile X, size of the CGG repeat, and genetic background. At present the ‘premutation’ class of allele can only be defined either empirically, by its observed progression over several generations to a full mutation or arbitrarily, by creating an artificial threshold of presumed ‘stability’. A most important part of future screening surveys will be to investigate the families of all subjects with higher than average numbers of repeats and to assess the stability of such borderline expansions through meiosis, while relating this to their haplotypes for linked markers. Thus it may eventually become possible to define the premutation state more rigorously in a single person.

We would like to thank the staff of the Community Health Department, Odstock and of the Department of Community Paediatrics, Isle of Wight for their help with records and assistance at clinics, which was invaluable.

Table 6 Comparison of FRAXAC1 genotype and CGG repeat number.

<table>
<thead>
<tr>
<th>CGG repeat number</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs</td>
<td>Exp</td>
<td>Obs</td>
<td>Exp</td>
</tr>
<tr>
<td>≤(26)</td>
<td>0</td>
<td>2·3</td>
<td>4</td>
<td>5·4</td>
</tr>
<tr>
<td>(27–33)</td>
<td>4</td>
<td>7·9</td>
<td>2·1</td>
<td>4·1</td>
</tr>
<tr>
<td>(≥34)</td>
<td>4</td>
<td>1·3</td>
<td>0·4</td>
<td>0·9</td>
</tr>
</tbody>
</table>

A D C Total

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>43</td>
<td>3</td>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>2·3</td>
<td>35·4</td>
<td>8·3</td>
<td>5·4</td>
<td>107</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1·3</td>
<td>7</td>
</tr>
<tr>
<td>0·4</td>
<td>5·4</td>
<td>1·3</td>
<td>1·3</td>
<td>100</td>
</tr>
</tbody>
</table>

Population studies of the fragile X: a molecular approach


Population studies of the fragile X: a molecular approach.

P A Jacobs, H Bullman, J Macpherson, S Youings, V Rooney, A Watson and N R Dennis

doi: 10.1136/jmg.30.6.454

Updated information and services can be found at:
http://jmg.bmj.com/content/30/6/454

These include:

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/